HARMONIC DISTORTION IN FREQUENCY MODULATION RECEPTION. PART I

By K. V. KRISHNA PRASAD

(*Received for publication, July 2, 1951*)

ABSTRACT. The two signal distortion theory of frequency modulation reception has been developed by a new method and a brief treatment of this theory togther with its extension to the more complicated case of n-signals is given.

INTRODUCTION

Transmission is possible by more than one path with the frequencies used in frequency modulation. So, two or more signals, having nearly the same amplitude, if induced in the antenna, will give rise to considerable distortion.

Large objects, such as hills or high buildings, reflect and absorb the waves and cause interference. This difficulty is encountered also in television reception. In fact, higher frequencies employed here tend to enhance this difficulty in as much as the phase changes encountered are great. This interference causes light and dark bands in the picture resulting in synchronization difficulties when frequency modulation is used on video signals. So, a theoretical and experimental study of multipath distortion would be very interesting in determining the factors which contribute to this type of distortion. The present paper is restricted only to a treatment of the theoretical aspects of the harmonic distortion in frequency modulation reception while a treatment of its experimental aspects is reserved for a latter one to follow.

The existing theories, such as those of Corrington (1945) and Meyers (1946) for the case of two signals, are complicated and they involve approximations from tables and charts in arriving at the final expression for the distortion term. In addition to these difficulties the results derived from them do not lend themselves to an easy experimental verification. The present development of the theory, treated here, involves the methods of contour-integration, leading to a rigorous mathematical treatment, which lends itself to physical interpretation. The general principles of the method are too well known to need any further elucidation. In recent years this method has been widely used in treating similar problems in the fields of electric transients and computer-electronics.

THEORY

1. *Case of two signals* :—Now to begin with, let a sinusoidal carrier frequency modulated by a single sinusoidal modulating frequency

be assumed for the transmitted wave. Consider an ideal limiter which is independent of the frequency of the input signal. Assume also a zero order characteristic to insure limiting at low amplitudes*. Noise is the controlling factor in any practical consideration where multipath interference reduces the strength of the signal at the receiver below it. This noise may, however, be neglected as an approximation. To develop the theory of two signal distortion, let the two signals be represented by :

$$e_1 = E_0 \sin y_1, \text{ the desired one}$$

and

$$e_2 = E_1 \sin y_2, \text{ the undesired one}$$

with

$$y_1 = \omega_c t + h \frac{\omega_d}{\omega_p} \sin \omega_p t$$

and

$$y_2 = \omega_c t + \beta + \cancel{h} \frac{\omega_d}{\omega_p} \sin (\omega_p t + \alpha) \qquad \ldots \quad (1)$$

where, $\omega_c = 2\pi \times$ carrier frequency

$\omega_p = 2\pi \times$ modulating frequency

$\omega_d = 2\pi \times$ maximum deviation frequency

$\alpha = \omega_p t_0$ and $\beta = \omega_c t_0$ with $t_0 =$ time delay in seconds

and, $h =$ factor proportional to the amplitude of signal input at the transmitter.

Then the resultant E_R of the two signals may be written as :

$$E_R = E \sin [(y_1 + y_2)/2 + \phi(t)] \qquad \ldots \quad (2)$$

where,

$$E = \sqrt{E_0^2 + E_1^2 + 2E_0E_1 \cos (y_1 - y_2)}$$

and

$$\tan \phi = \frac{E_0 - E_1}{E_0 + E_1} \tan (y_1 - y_2)/2$$

Assuming the signal above limiter level and E_0, E_1 constant with respect to time, the discriminator output E is given by

$$E = \frac{d}{dt}\left[\frac{y_1 + y_2}{2} + \phi(t)\right] \qquad \ldots \quad (3)$$

$$= \omega_c h \frac{\omega_d}{2} \cos \omega_p t + h \frac{\omega_d}{2} \cos (\omega_p t + \alpha) + d\phi/dt \qquad \ldots \quad (4)$$

The last term in equation (4) *viz*.

$$\frac{d\phi}{dt} = \frac{E_0^2 - E_1^2}{2E_0E_1} \cdot \frac{d\xi/dt}{\frac{E_0^2 + E_1^2}{2E_0E_1} + \cos 2\xi} \qquad \ldots \quad (5)$$

where, $\xi = \frac{1}{2}[h(\omega_d/\omega_p) \sin \omega_p t - h(\omega_d/\omega_p) \sin (\omega_p t + \alpha) - \beta]$ $\qquad \ldots \quad (6)$

* This fact has been borne out in the experimental tests by operating above the limiter level.

Now $f(\xi)=1/(\sigma+\cos 2\xi)$ being a symmetrical function may be expanded into the series :

$$f(\xi)=a_0/2+\sum_{\nu=1}^{\infty}(a_\nu \cos \nu\xi+b_\nu \sin \nu\xi) \quad \ldots \quad (7)$$

where the Fourier coefficients,

$a_\nu=0$ for odd values of ν

and $\quad b_\nu=0$ for all values of ν

Hence,

$$a_{2N}=\frac{1}{\pi}\int_{-\pi}^{\pi}\cos 2N\xi/(\sigma+\cos 2\xi)\,d\xi \quad \ldots \quad (8)$$

which by contour integration and substitution of $Z=e^{2i\xi}$ becomes :

$$a_{2N}=\frac{1}{2\pi i}\int\frac{Z^N+Z^{-N}}{1+2\sigma Z+Z^2}\,dZ \quad \ldots \quad (9)$$

The path of integration is the unit circle $|Z|=1$ and the poles are $Z_0=0$, $Z_1=-E_1/E_0$ and $Z_2=-E_0/E_1$

Since $E_0>E_1$, consider only the poles $Z_0=0$ and $Z_1=-E_1/E_0$, their respective residues being $(Z_1^N-Z_2^N)/(Z_1-Z_2)$ and $(Z_1^N+Z_2^N)/(Z_1-Z_2)$, so that :

$$a_{2N}=4E_1E_0/(E_0^2-E_1^2)\ (-r)^N, \text{ where } r=E_1/E_0 \quad \ldots \quad (10)$$

and hence

$$\frac{d\phi}{dt}=[h\omega_d \cos \omega_p t-h\omega_d \cos(\omega_p t+\alpha)].[\tfrac{1}{2}+\sum_{N=1}^{\infty}(-r)^N \cos 2N\xi] \quad \ldots \quad (11)$$

Substituting the value of $d\phi/dt$ given in expression (11) in equation (4), the expression for the discriminatoı output E becomes :

$$E=\omega_c+h\omega_d \cos \omega_p t-[2h\omega_d \sin \alpha/2 \sin(\omega_p t+\alpha/2)]$$
$$\times\left[\sum_{N=1}^{\infty}(-r)^N \cos 2N\xi\right] \quad \ldots \quad (12)$$

from which the instantaneous frequency f_i may be computed as

$$f_i=E/2\pi=f_c+hf_d \cos \omega_p t-[2hf_d \sin \alpha/2 \sin(\omega_p t+\alpha/2)]$$
$$\times\sum_{N=1}^{\infty}(-r)^N \cos 2N\xi \quad \ldots \quad (13)$$

The distoıtion D represented by the last term of equation (13) is given by

$$D=lf_p \sin(\omega_p t+\alpha/2)\sum_{N=1}^{\infty}(-r)^N \cos(2N\xi)$$

or in the expanded form,

$$D=lf_p \sin(\omega_p t+\alpha/2)\sum_{N=1}^{\infty}(-r)^N \cos[Nl \cos(\omega_p t+\alpha/2)+N\beta] \quad \ldots \quad (14)$$

where $l=2hf_d/f_p \sin \alpha/2$.

Equation (14) obviously can be expanded into a series of Bessel functions and hence can be written in the form :

$$D=2f_p\left\{\sum_{N=1}^{\infty}\frac{(-r)^N}{N}\sin N\beta\sum_{m=1}^{\infty}(-1)^m(2m)J_{2m}(Nl)\sin 2m\theta\right.$$

$$\left.+\sum_{N=1}^{\infty}\frac{(-r)^N}{N}\cos N\beta\sum_{m=0}^{\infty}(-1)^m(2m+1)J_{1+2m}(Nl)\sin(2m+1)\theta\right\}\quad\ldots\quad(15)$$

where $\theta=\omega_p t+\alpha/2$.

For purposes of experimental verification, this expression can be further simplified and put into a more convenient form by the substitution

$$J_m(l)=\sum_{p=0}^{\infty}\frac{(-1)^p(l/2)^{m+2p}}{p!(m+p)!}$$

whence :

$$D=f_p\left[\sum_{p=0}^{\infty}\sum_{m=0}^{\infty}2m/p!(m+p)!\,(\sin m\theta)(l/2)^{m+2p-1}\frac{\partial^{m+2p-1}R}{\partial\beta^{m+2p-1}}\right]\quad\ldots\quad(15a)$$

where $R=\sum_{n=1}^{\infty}(-r)^N(\cos N\beta)l/2$ for the sake of brevity.

The parameters which are significant in equation (15) or (15*a*) are r, the ratio of the undesired to the desired signal, l, the product of h, the input signal strength and f_d, the deviation frequency, f_p, the modulating frequency, m, the order of the harmonic and finally α and β, the factors involving the time delay t_0. From the stand point of high distortion, the case of interest is when r is very near unity but not exactly unity. The distortion in this case is rigorously represented by equation (15 or (15*a*). The distortion is zero when l is large ; that is, for a system having high input signal and high deviation frequency. Further, the distortion is zero when $\alpha=2\pi N$ and the even order modulations vanish for $\beta=N\pi$. Lastly, the distortion increases with the modulating frequency f_p. So from the point of view of high distortion, the case of low signal amplitudes is interesting when h and hence $l=2hf_d/f_p\sin\alpha/2$ is small. Then only the first term of the Bessel series expansion in equation (15) is important. So by putting $p=0$ in equation (15*a*) the resulting simplified expression for the distortion is :

$$D=f_p\left|\sum_{=1}\frac{2\sin m\theta}{(m-1)!}(l/2)^{m-1}\frac{\partial^{m-1}R}{\partial\beta^{m-1}}\right|\qquad(16)$$

From the expression it is easy to compare the amount of various harmonics by letting m have the run of values $m=1,2,3,\ldots$ and the results are summarized in Table I below. In these expessions $u=(r^2+r\cos\beta)/(1+2r\cos\beta+r^2.)$

The results given in Table I will be dealt with in detail in part II but suffice it to say that the nth-harmonic distortion is directly proportional to the $(n-1)$ th power of the deviation frequency f_d, $(n-1)$th power of the input amplitude h, and the modulating frequency f_p for small values $\alpha=2\pi f_p t_0$

TABLE I

Order of the harmonic m	Simplified expressions for the distortion (D)
Fundamental $m=1$	$D_1 = hf_d \cos \omega_p t - lf_p \sin (\omega_p t + \alpha/2) \cdot u$
Second $m=2$	$D_2 = -2hf_d/f_p \, \partial u/\partial\beta \sin^2 \alpha/2 / \sqrt{1+4u^2 \sin^2 \alpha/2}$
Third $m=3$	$D_3 = h^2 (f_d/f_p)^2 \partial^2 u/\partial\beta^2 \sin^3 \alpha/2 / \sqrt{1+4u^2 \sin^2 \alpha/2}$
Fourth $m=4$	$D_4 = h^3/3 (f_d/f_p)^3 \partial^3 u/\partial\beta^3 \sin^4 \alpha/2 / \sqrt{1+4u^2 \sin^2 \alpha/2}$
nth $m=n$	$D_n = 2h^{n-1}/(n-1)! \, \partial^{n-1} u/\partial\beta^{n-1} (f_d/f_p)^{n-1} \sin^n \alpha/2 / \sqrt{1+4u^2 \sin^2 \alpha/2}$

where t_0 is the time delay ; and so in any practical consideration it is advisable to operate at high deviation frequency and high modulating frequency if any appreciable distortion is desired.

2. *Case of n-signals*:—A preliminary exprimental survey has shown that most of the distortion is caused by the combination of two signals, the desired and undesired ones of very nearly of the same amplitude. Nevertheless, it will be interesting to study the arrival of more than two signals at the receiver by multiple reflections. So a brief treatment of the general case of n-signals is worth considering from the theoretical point view.

Let the desired signal and the n-interfering signals be represented by

$$\begin{aligned} e_1 &= E_0 \sin [\omega_c t + hf_d/f_p \sin \omega_p t] \\ e_2 &= E_1 \sin [\omega_c t + \beta_1 + hf_d/f_p \sin (\omega_p t + \xi_1)] \\ & \\ e_n &= E_n \sin [\omega_c t + \beta_n + hf_d/f_p \sin (\omega_p t + \xi_n)] \end{aligned} \tag{17}$$

where $\beta_n = \omega_p t_n$ and $\xi_n = \omega_c t_n$ with $n = 1, 2, 3, \ldots n$

The resultant E_R of these signals is given by :

$$E_R = \sum_{j=0}^{n} E_j \sin (\omega_c t + \zeta_j)$$

where $\zeta_j = \beta_j + hf_d/f_p \sin (\omega_p t + \xi_j)$

On expanding,

$$\begin{aligned} E_R &= [\sum E_j \cos \zeta_j] \sin \omega_c t + [\sum E_j \sin \zeta_j] \cos \omega_c t \\ &= \sqrt{[\sum E_j \cos \zeta_j]^2 + [\sum E_j \sin \zeta_j]^2} \sin (\omega_c t + \tau) \end{aligned}$$

where $\tau = \text{arc tan} \sum_{j=0}^{n} E_j \sin \zeta_j / \sum_{j=0}^{n} E_j \cos \zeta_j$

By putting

$$\zeta_1 = \zeta_0 + \epsilon_1$$
$$\zeta_2 = \zeta_0 + \epsilon_2$$

where $\zeta_0 = hf_d/f_p \sin (\omega_p t)$

the expression for τ could be simplified to :

$$\left.\begin{aligned}\tau &= \zeta_0 + \tan^{-1}\frac{\sum_{j=1}^{n} E_j \sin\epsilon_j}{E_0 + \sum_{j=1}^{n} E_j \cos\epsilon_j} \\ &= \zeta_0 + \psi \\ \text{where}\quad \psi &= \tan^{-1}\frac{\sum_{j=1}^{n} E_j \sin\epsilon_j}{E_0 + \sum_{j=1}^{n} E_j \cos\epsilon_j}\end{aligned}\right\} \qquad (18)$$

Whence,

$$E_R = \sqrt{[\Sigma E_j \cos\zeta_j]^2 + [\Sigma E_j \sin\zeta_j]^2}\ \sin(\omega_c t + hf_d/f_p \sin\omega_p t + \psi) \qquad (19)$$

Hence the distortion is given by :

$$D = \frac{d}{dt}\tan\psi^{-1}$$

$$= \frac{d}{dt}\tan^{-1}\frac{\sum_{j=1}^{n} E_j/E_0 \sin\epsilon_j}{1 + \sum_{j=1}^{n} E_j/E_0 \cos\epsilon_j}$$

which on analogy with the two signal theory* can be represented by :

$$D = -\sum_{s=1}^{\infty}[f(\epsilon_j)]^s \sin\sum_{p=1}^{n}(E_j/E_0)_p \sin(p\epsilon_j + \eta_p) \qquad \ldots \quad (20)$$

where $f(\epsilon_j)$, containing the amplitudes of the harmonics is a complicated function of ϵ_j The problem becomes then that of expanding a sine of a Fourier series into a suitable form, let alone the other complicated parameters involved. The method of approach is as follows :

The rather well known result in the theory of Bessel functions viz,

$$e^{ic_p \sin(p\theta + \delta_0')} \sin p\theta + \delta_0 = \sum_{p=-\infty}^{\infty} e^{iq(p\theta + \delta_p)}\xi_q(cp)$$

is generalized by replacing $c_p \sin(p\theta + \delta_p)$ on the left hand side by the expression $\phi_n = \sum_{p=1}^{n} c_p \sin(p\theta + \delta_p)$ whence

$$e^{i\phi_n} = \prod_{p=1}^{n} e^{ic_p \sin(p\theta + \delta_p)}$$

$$= \prod_{p=1}^{n}\left\{\sum_{q_p=-\infty} e^{iq_p\delta_p}\,\xi_{q_p}(c_p)e^{ipq_p\theta}\right\}$$

$$= \sum_{q_n=-\infty}^{\infty}\ \sum_{q_{n-1}=-\infty}^{\infty}\cdots\sum_{q_1=-\infty}^{\infty} e^{i\sum_{p=1}^{n} q_p\delta_p}\left\{\prod_{p=1}^{n}\xi_{q_p}(c_p)\right\}e^{i\sum_{p=1}^{n} pq_p\theta}$$

$$= \sum_{m=-\infty}^{\infty}\Gamma_m e^{im\theta}$$

where the Γ_ms are complex, and they may be evaluated by restricting the summation over the q's such that $\sum_{p=1}^{n} pq_q = m$.

*c.f., equation (14)

Hence,
$$\Gamma_m = \sum_{(q_p)} \left[e^{i\sum q_p \delta_p} \left\{ \prod J_{q_p}(e_p) \right\} \right.$$

which can be re-written so as to conform to the notation used in equation (20) as :

$$\Gamma_m = \sum_{(q_p)} \left[e^{i\sum_{1}^{n} q_p \eta_p} \left\{ \prod_{p=1}^{n} J_{q_p} (E_j / E_0)_p \right\} \right] \tag{21}$$

The m th coefficient of the sine of the Fourier series is easily evaluated by taking the imaginary part of $\Gamma_m e^{im\epsilon_j} + \Gamma_{-m} e^{im\epsilon_j}$. Tables are given by Strachey and Wallis (1946) for evaluating some of these parameters.

Thus, while in the previous papers of Corrington (1945) and Meyers (1946), the final expressions for distortion for the case of two signals are presented in a complicated form involving too many parameters without any attempt being made to simplify the results, the present paper gives the final result in a useful series form enabling easy computation of the various harmonic contents for a particular case under test and this will be explained in detail together with experimental results, in part II.

The case of n-signals* (Krishna Prasad, 1951) presents a rather unique and complex situation of resolving the sine of a Fourier series and a method of approach is suggested: The results, however, cannot be put into a useful practical form on account of the highly complicated nature of $f(\epsilon_j)$ in equation (20).

ACKNOWLEDGMENT

This work forms part of the the thesis submitted by the author to the California Institute of Technology, Pasadena, U. S. A. for the Degree of Electrical Engineer under the title, Distortion in Frequency Modulation Reception'.

NATIONAL PHYSICAL LABORATORY OF INDIA,
NEW DELHI.

REFERENCES

Corrington, M. S , 1945, *Proc. Inst. Rad. Engrs* , **33**. 878
Krishna Prasad, K. V., 1951, *Science & Culture* , **17**, in Press.
Meyers, S. T., 1946, *Proc. Inst. Rad. Engrs.*, **34**. 256.
Strachey, C., and Walis, P. J., 1946, *Phil. Mag.*,**2**, 84

REVIEW

(4)

Static and Dynamic Electricity—By William R. Smythe. Pp xxi+616. McGraw-Hill Book Company, New York, Toronto, London, 1950. Price $ 8.5.

This is the second edition of the book, the first edition of which was published in 1939. As pointed out in the preface to the first edition, special stress has been laid on solution of problems in electricity and magnetism which are generally encountered in research. A basic knowledge in vector analysis has been assumed to be possessed by the student in treating the problems, but the use of contour integration has been avoided.

In electrostatics, method of images, Lengendre polynomials and Bessel's, functions have been applied to problems on potential. In electrodynamics, besides chapters on Flow of current in networks and two and three dimensional conductors, Eddy currents, Magnetism, Electromagnetic waves, etc., there are important chapters on Electromagnetic radiation and Wave guides and Cavity resonators. There is also a chapter on specal relativity and the motion of charged particles, The mks (metre-kilogram second) system of units has been used throughout, but in an Appendix the factors of conversion to cgs units have been given in different tables.

The book will be found useful to students studying for the B. Sc. (Hons.) and M. Sc. degrees of Indian Universities. It is also a good reference book which can be used profitably by research workers.

The paper used and quality of printing leave nothing to be desired.

S. C. S.